**Mock Test 1 - Big Data Question no. 9**

**9.** Calculate the average price of all products in each category and display the results along with the category name.

**Sol.**

from pyspark.sql import SparkSession

from pyspark.sql.functions import avg

# Create SparkSession

spark = SparkSession.builder \

.appName("AveragePriceByCategory") \

.getOrCreate()

# Assuming you already have a DataFrame named "products" containing the "Products" table

# Calculate the average price of products in each category

average\_price\_df = products.groupBy("category") \

.agg(avg("price").alias("average\_price"))

# Show the result

average\_price\_df.show()

# Stop the SparkSession

spark.stop()